LEVEL

ADA056499

UNITED STATES AIR FORCE
AIR UNIVERSITY
AIR FORCE INSTITUTE OF TECHNOLOGY
Wright-Patterson Air Force Base, Ohio

78 07 07 022

AFIT/GSM/SM-77D-20

1 LEVEL

AD A056499

9 Master's thesis,

11 Dec 77

12 94 p.

DDC
RECEIVED
JUL 21 1978
B

6 A JUDGMENT ANALYSIS APPROACH TO EXAMINING A "WHOLE PERSON" CONCEPT OF AIR FORCE PROMOTIONS,

THESIS

14 AFIT/GSM/SM-77D-20

10 Philip E. Glenn
Capt USAF

AD No.
DDC FILE COPY

Approved for public release; distribution unlimited

78 07 07 022
012 225

AFIT/GSM/SM-77D-20

A JUDGMENT ANALYSIS APPROACH TO EXAMINING A "WHOLE PERSON" CONCEPT OF AIR FORCE PROMOTIONS

THESIS

Presented to the Faculty of the School of Engineering
of the Air Force Institute of Technology
Air University
in Partial Fulfillment of the
Requirements for the Degree of
Master of Science

by

Philip E. Glenn, B.S.
Capt USAF
Graduate Systems Management

December 1977

Approved for public release; distribution unlimited.

Preface

This research effort would not have been possible without the help and guidance of many individuals. I wish to express my sincere appreciation to Lieutenant Colonel Adrian M. Harrell who sparked my initial interest in this topic. As the thesis advisor, Colonel Harrell offered valuable guidance and constructive criticism throughout the research effort. His undaunted enthusiasm was always an inspiration to me.

In addition, I owe a debt of gratitude to Major Charles W. McNichols who not only served as the thesis reader, but also provided me with timely and efficient expertise in computer programming techniques and statistical analysis.

Numerous personnel assigned to the various professional military education schools at Maxwell Air Force Base, Alabama contributed their valuable time and effort in order to assist in the distribution and collection of the research data collection instrument. In particular, I would like to thank Lieutenant Colonel Baker, AU/EDV, Dr. Ray Maize, AWC/EDR, Major Ned Nelson, ACSC/EDV, and Captain Jim Simpson, SOS/EDV.

The final copy of this thesis would not have been possible without the professional editing and typing abilities of Miss Joyce Wilson. I wish to thank her for the contributions she made to this research.

Most of all, I sincerely thank my wife, Sylvia, for the numerous times she read and corrected the many draft copies of this thesis. However, more importantly, I wish to thank her for the love, understanding, and sympathy she shared with me during the long hours which this thesis demanded.

Philip E. Glenn

List of Figures

List of Tables

Contents

Page

AFIT/GSM/SM-77D-20

Abstract

The United States Air Force stresses the "whole person" concept in the evaluation of all commissioned officers who are eligible for promotion. However, no specific policy has been established by the Air Force for applying this whole person evaluation. The separate promotion boards are presented a whole person chart which recommends factors and areas for evaluation. The criteria depicted on this chart, however, are not mandatory nor do they have precedence over any other criteria. It appears, then, that each promotion board establishes the policy which is utilized in selecting commissioned officers for promotion.

This research effort investigates the extent to which a whole person concept of promotion evaluation influences the beliefs held by Air Force officers concerning the promotability of captains to the rank of major. In order to do so, a judgment analysis approach known as policy capturing was utilized.

A random sample of commissioned officers attending Squadron Officers' School, Air Command and Staff College, and Air War College was administered a decision-making exercise. In this exercise, the selected officers were asked to evaluate the promotability of 32 hypothetical captains based upon five key promotion factors which were used as decision criteria. The five promotion factors used in the research were: 1) Officer Effectiveness Report ratings, 2) professional military education, 3) formal education, 4) assignment history, and 5) aeronautical rating. Data collected from this decision exercise were used to test specific hypotheses concerning the decision-making behavior of the selected officers.

The findings of the research indicate that Air Force officers are aware of the whole person concept of promotion evaluation and utilize it in their promotability decisions. Junior officers, however, placed significantly different weights on three of the five factors when compared with the weights placed on the same factors by senior officers. The data indicates that officers combine the decision criteria in an essentially linear fashion when rendering their promotability decisions. Finally, the findings showed that the individual decision-making policies among the officers were not homogeneous, which indicates that there is a wide variety in the way Air Force officers make promotability decisions.

A JUDGMENT ANALYSIS APPROACH TO
EXAMINING A "WHOLE PERSON" CONCEPT OF
AIR FORCE PROMOTIONS

I Introduction

"Best qualified" is a difficult term that mystifies hundreds of Air Force officers each year. It represents the cut-off point or the pass-fail mark in field grade promotion competition. Those chosen for advancement are the individuals the Air Force calls "best qualified." For the several hundred officers who are not selected, the inevitable question in their minds is, How does the Air Force define "best qualified"? To answer this question, it would be helpful to develop some background information on the promotion process.

The Air Force Promotion Process

The authority for promoting commissioned officers can be found in the United States Code, Title 10 (USC 10). In general terms, this statute allows the Secretary of the Air Force to direct a promotion board to recommend a specified number of officers for promotion. An officer may be recommended by one of two different promotion systems: 1) fully qualified status or 2) best qualified status. Both systems are cited in public law as alternative procedures available to promotion boards in considering officers for advancement.

The fully qualified selection process is utilized when there are no statutory limitations on the number of active duty commissioned officers in a particular grade. The best qualified selection method is utilized

when there are limitations placed on the number of officers in a specified grade (USC 10, 1970:1,956). Consequently, the best qualified method has been utilized for the past several years because of grade limitations established by law.

Under the best qualified selection system, quotas are set which are smaller than the total number of officers being considered for promotion. These quotas represent the highest number of competitors each board may pick for advancement. Those chosen are identified as "best qualified." However, USC 10 does not prescribe the specific criteria for selecting officers under either the fully qualified or best qualified methods.

Likewise, the Air Force regulations relating to the promotion of commissioned officers do not specify criteria to be used for promotion under either system (AFR 36-11, 1974, and AFR 36-89, 1975). These regulations basically mirror the intent of public law as prescribed in USC 10. Therefore, under either the fully qualified or best qualified system, selection is made on the basis of the judgment of board members of the relative qualifications of the officers under consideration. How these judgments are reached, and/or on what basis they are formed, amounts to the "why" of the selection process.

Specific information about this aspect of the promotion program is very difficult to acquire. The Air Force is very secretive about the selection process. Board proceedings are kept under tight controls. All that generally surfaces are the names of selectees and some summary statistical data on the individuals who were selected.

Some information for this research effort was provided by the Directorate of Personnel Program Actions, Headquarters Air Force Military Personnel Center. In a cover letter dated June 13, 1977 the Chief, Officer Promotions and Regular Appointments Branch stated that "The

Secretary of the Air Force did not provide information to temporary major boards." Board members were given a letter of instruction which provided general information for the conduct of the board. For example, in a copy of such a letter it was specified what percentage of the officers to be considered could be recommended for promotion. No specific or quantifiable criteria were given in this letter (Letter of Instruction, 1976).

In addition to the letter of instruction, the board members were presented a "whole person concept" chart; the format of which is depicted in Fig. 1.1.

Total Evaluation

Factors	Evaluate
Performance	OERs
Education	Level/Utilization
Breadth of Experience	Where/When/What
Job Responsibility	Scope/Exposure
Professional Competence	Expertise of Specialist
Combat/Achievements	Awards/Decorations
Leadership	Staff Command

Other Factors/Evaluations May Apply

Fig. 1.1. Whole Person Concept (Source: Hq AFMPC)

It would appear that the "whole person" chart is the primary source of evaluation criteria for determining the "best qualified" officers. However, it should be noted that these criteria are not quantified or mandatory nor are any of them given precedence over any other with perhaps the exception of the Officer Effectiveness Reports (OER). Apparently the importance to be attached to the criteria in the guidelines is left to the discretion of the board member. Moreover, having taken an oath to base their selections on the "...special fitness of the officers and the efficiency of the United States Air Force" (Letter of Instruction, 1976),

board members could well feel free to establish their own criteria as to qualities an officer should have to enhance the efficiency of the Air Force. This could lead to unequal evaluation of officers based on the personal philosophy of different board members. Furthermore, this lack of specific or binding criteria could lead to the use of inconsistent standards by the several selection boards (Ford, 1975). Even so, the Air Force has never specified precisely what was required of those under consideration to be selected as "best qualified" among their peers (Foster, 1973).

Statement of the Research Problem

The Air Force stresses the importance of the whole person concept in the evaluation of all officers eligible for promotion. However, the policy for applying this total evaluation is determined by each individual promotion board. This research effort investigates the extent to which the whole person concept influences the beliefs held by Air Force officers concerning the promotability of captains to the rank of major. In order to do so, the judgment policies of Air Force officers regarding this decision were identified using a research approach known as policy capturing.

Identifying Individual Judgment Policies

The best and possibly the only method to accurately and objectively identify judgment policies is through an empirical analysis of actual judgments. Such a process has come to be known as "policy capturing" (Taylor, 1973). It is, in simple terms, a mathematical description of judgment policy which can be used to predict and understand future judgments. Policy capturing analyzes decisions and yields a mathematical model of the cues on which decisions are based, weighting each according

to the influence on the decision. This procedure is accomplished through the use of various mathematical models which will be described in the next chapter.

This research effort utilizes policy capturing to investigate the judgment policies of a random sample of Air Force officers. This was accomplished through the use of a decision-making exercise which contained specific information relating to five factors of promotion. Utilizing these five factors of promotion as decision criteria, the sampled officers were asked to evaluate or judge the promotability of a group of 32 hypothetical Air Force captains. The five promotion factors used in the research were: 1) Officer Effectiveness Report ratings, 2) professional military education, 3) aeronautical rating, 4) assignment history, and 5) formal education. Support for the use of these five promotion factors is discussed in Chapter III.

In addition, it is important to emphasize that the decision exercise was specifically designed to elicit the whole person evaluation. This point is also discussed further in Chapter III.

Objectives of the Research

The fundamental objective of this research effort was to collect empirical data for investigating the individual judgment policies of Air Force officers regarding promotability decisions. This data was used to examine the following research hypotheses:

Primary Hypothesis

H1: Officers incorporate a whole person concept of ev·luation in their promotability decisions by utilizing all the selected promotion factors in the promotability decision process.

Additional Hypotheses

H2: Junior officers place the same relative weights upon the criteria used for their promotability decisions as do senior officers.

H3: Air Force officers combine the promotability decision criteria in an essentially linear fashion to render a judgment.

H4: Individual officers accurately specify the relative weights they place upon the criteria used to render their promotability decisions.

H5: Individual judgment policies are homogeneous within a group of officers, indicating that Air Force officers, as a group, make decisions relating to promotions in essentially the same manner.

Scope and Limitations of the Study

Because of the nature of this study, one might conclude that it is an attempt to simulate the promotion board process. This is not the case. The study attempts to define how individuals render promotability decisions. In the board process, intragroup communication is an important part of determining the selection policy. By design, this study did not allow either intragroup or intergroup communication in order to identify the individual policies of the selected officers.

To simplify the decision-making exercise which was used as the source of data, only five factors relating to promotion potential were chosen. These five factors are not necessarily the ones used in actual selection processes, although some effort was made to utilize realistic factors. It is realized that other factors may come into play when determining the promotion potential of an individual. The goal of the exercise was to make the decisions relatively simple and, yet, not detract from the realism of the decision.

II Modeling Human Judgment

Interest in the area of human judgment has probably been in existence for almost as long as judgment-making has been in existence. Most attempts at defining this ill-defined process of human judgment have been via the use of more clearly defined mathematical models. This area of analysis is found under various titles but is most often referred to as Policy Capturing, a label which originated at the Air Force Human Resources Laboratory, Lackland Air Force Base, San Antonio, Texas (Jones, et al., 1976:1).

As with any relatively complex concept, the technical terms found in the literature sometimes cloud the overall objective. So, before proceeding further, some definitions are in order.

Definition of Technical Terms

The following definitions will hopefully clarify the terminology used in the judgment modeling literature:

Policy Capturing (or judgment modeling) - "...the building of a model which, given the same information the individual has, will accurately reproduce his judgments based on that information" (Smith, 1972).

Policies - Management constraints delineating feasible alternatives for decision-making.

Decision(s) - Specific recommendation for implementation to achieve goal within policy. Decision-making usually involves a specific action which influences the near future.

Cue - External stimulus or information which is used for arriving at a decision.

Configural cue utilization - A patterned combination of external stimuli or information that cannot be modeled by a first-order (linear) mathematical equation.

Paradigm - An example, model, or pattern.

The Case for Modeling Human Judgment

Many people believe that human judgment (decision-making) is too complicated to predict with any real accuracy. First of all, opponents feel that the utilization of external stimuli, or cues, to arrive at decisions seems to be dependent upon numerous environmental, educational, and experimental behavior elements learned over time by every individual. Secondly, judgment also seems to involve physiological dynamics, such as the state of health of an individual, the diet of the individual, and other chemical and hormonal balances and imbalances. Finally, the human judgment process seems to incorporate varying degress of intuition (O'Berry, 1977:20).

Despite the apparent complexities involved with modeling human judgment, there are many researchers who believe that the manner in which humans use available information in making decisions can be modeled. Paul J. Hoffman feels that if a mathematical model effectively predicts judgments for any given set of information, then the judgment process has been adequately described (Hoffman, 1960:117).

Studies have shown that, although mathematical models based on quantitative analyses may not be optimal, the consistent application of these models often leads to decisions that are superior to those of the individuals who are being modeled. This arises from the fact that humans tend to be erratic in their judgments. This in turn generates error that reduces the accuracy of the decisions. The model filters out this error and is, therefore, able to outperform the decision-maker whose judgment it was designed to simulate (Slovic, *et al.*, 1972).

The Framework for Judgment Models

The majority of work in the analysis of human judgment has been within two frameworks: the Bayesian and the regression. The Bayesian approach is founded on the elementary theorem of statistics known as Bayes' Theorem (developed in 1763 by the Reverend Thomas Bayes). Since the method of Bayesian analysis was not utilized in this study, further explanation of this methodology will not be presented. The interested reader is referred to the excellent review article by Slovic and Lichtenstein (1971) for a detailed treatment of the subject.

Research within the regression approach has primarily centered on two paradigms: the correlational and the functional measurement. The correlational model is primarily concerned with using correlational statistics to relate the decision of a judge (or judges) to the information which resulted in that decision.

The functional measurement model is concerned not only with correlating the information and the decision but also with analyzing the relationship of the judge to his environment. In particular, this approach studies how well a decision-maker can learn to perform in an experimental situation. As a consequence, the correlational paradigm is concerned with providing descriptive models, whereas the functional measurement paradigm is more concerned with the task of learning (Jones, *et al.*, 1976).

The model used in this study was the linear model suggested by Hoffman (1960) which is based, in turn, upon the lens model presented by Brunswik (1952). As used in this study, the model falls under the regression framework and utilizes the descriptive qualities of the correlational paradigm. The model is founded upon the notion that a judge's decisions represent linear combinations of the available cues.

The Multivariate Linear Regression Model

The work of Brunswik, Hoffman, and many others in the field of modeling human judgment is based upon the process of fitting data to a linear equation through the method of least squares regression. The resulting equation is referred to as a regression model. The following discussion of regression analysis is provided in order to clarify the model utilized in this study. Some statistical background on the part of the reader is assumed.

The model

$$Y_i = B_0 + B_1X_{i1} + B_2X_{i2} + \varepsilon_i \qquad (1)$$

is called a first-order model with two independent variables. A first-order equation is linear in parameters and linear in the independent variables. Y_i is sometimes denoted as the criterion variable and the independent variables X_1 and X_2 are denoted as predictor variables. The parameters of the model are B_0, B_1, and B_2. The error term, ε_i, accounts for any variance in Y_i not explained by the predictor variables. If one were to plot the model described by Eq (1), the result would be a plane as shown in Fig. 2.1.

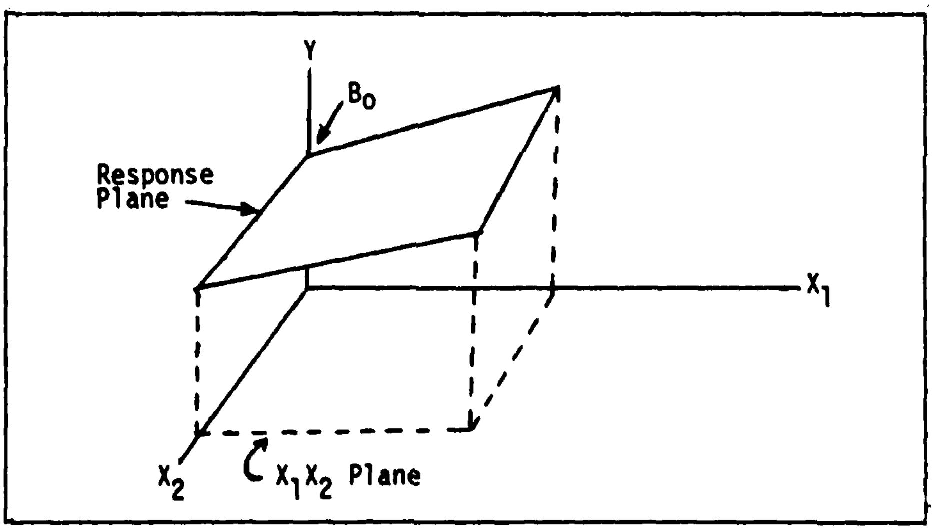

Fig. 2.1. Graph of Eq (1)

B_0 is the Y-intercept of the regression plane. If the scope of the model includes $X_1 = 0$, $X_2 = 0$, B_0 gives the mean response at $X_1 = 0$ and $X_2 = 0$. Otherwise, B_0 does not have any particular meaning as a separate term in the regression model. The parameter B_1 indicates the change in mean response per unit increase in X_1 when X_2 is held constant. Likewise, B_2 indicates the change in the mean response per unit increase in X_2 when X_1 is held constant.

When the effect of X_1 on the mean response does not depend on the level of X_2 and, correspondingly, the effect of X_2 does not depend on the level of X_1, the two independent variables are said to have additive effects or they do not interact. Thus, the first-order model is designed for independent variables whose effects on the mean response are additive or do not interact.

An example may clarify the model further. Suppose it is desirable to model the process of evaluating applicants for graduate school based upon the applicants' scores on Graduate Record Examinations (X_1) and undergraduate gradepoint average (X_2). Suppose further that the applicants will be evaluated by scoring the individual applicant on a scale from one (low potential for graduate work) to ten (high potential for graduate work). This evaluation score then can be defined as Y_i in the regression model, Eq (1).

In practice it is not usually feasible to examine all possible occurences of Y and X; therefore, statistical sampling is used to derive estimates of the actual parameters (B_0, B_1, B_2) of the regression model. Once these paramaters have been determined, the model then describes or predicts the evaluation score (from 1 to 10) based upon the Graduate Record Examination score (X_1) and the undergraduate gradepoint average (X_2).

If one were to calculate the response (Y_i) for every value of X_1 and X_2, the resulting graph would be a plane as shown in Fig. 2.1, page 10.

The simplistic model explained above would probably not be appropriate for evaluating graduate student applicants. After all, there are many other factors that figure into this determination. In order to measure how well a regression model describes the situation being modeled, statistical analysis derives a measurement denoted as R^2, the squared multiple correlation coefficient. This coefficient indicates the proportion of variance in the decision, Y, which is explained by the regression model.

The parameters B_1 and B_2 are frequently called partial regression coefficients because they reflect the partial effect of one independent variable when the other independent variable is included in the model and is held constant.

Standardized regression coefficients may be used to facilitate comparison between regression coefficients. Ordinarily it is difficult to compare regression coefficients because of differences in the units involved. To alleviate this problem, the regression coefficients are "standardized" which in turn makes them dimensionless. These standardized coefficients are sometimes referred to as beta coefficients and are denoted b_{iy}. The coefficient b_{iy} reflects the change in mean response for the ith trial (in units of standard deviations of Y) per unit change in the independent variable X_i (in units of standard deviations of X_i), when all other independent variables are held constant. The standardized regression coefficients show the effect of the given independent variable in the context of the other independent variables in the model. Changing the other independent variable in the model will usually change the

standardized regression coefficient when the independent variables are correlated among themselves. Hence, it is ordinarily not wise to interpret a standardized regression coefficient as reflecting the importance of the independent variable.

However, as will be seen later in Chapter III, the policy capturing technique employed for this study involves the use of orthogonal (independent) predictors in the regression equation. In this special case, a relative measure of the importance of the independent variable can be tied to the standardized regression coefficients.

Linear versus Configural Models

Linear models are frequently used in situations in which decisions are made on the basis of multiple codable inputs. These models sometimes are used normatively (*to aid the decision-maker*), sometimes contrasted with the decision-maker, sometimes used "paramorphically" (representing the decision-maker), and sometimes used to "bootstrap" the decision-maker (replacing the decision-maker with his representation) (Dawes and Corrigan, 1974:95). However, from the time Hoffman (1960) first introduced the linear regression model for modeling judgment, there has been debate between researchers favoring linear models versus those who favor higher order and/or configural (interaction) models.

The simplicity of the linear models disturbs a number of researchers. Many feel that the human judgment process is more complex and, therefore, requires models that account for interaction between cues. However, Lewis Goldberg (1971) compared the linear model to four other nonlinear models and found that the linear model was superior. By utilizing a clinical judgment experiment, Goldberg looked at a conjunctive model

(product of the cues), a disjunctive model (an inverse function of the cues), a logrithmic model (logrithm of the cues), and an exponential model (exponential function of the cues).

Hammond and Summers (1965) cite more than a dozen studies of classical or quasi-classical clinical judgments in which the accuracy of prediction derived from linear regression analysis was sufficiently great to suggest that judges are primarily linear in their mode of combining cues.

Bert F. Green, Jr., of Carnegie-Mellon University, feels that the use of standard configural techniques are essentially fishing expeditions. He states:

> The experimenter will covet any configural effect, any interaction that he can find. He cannot begin to examine all the possible nonlinear effects, and is very likely to miss those that are present, unless he knows where to fish (Kleinmuntz, 1968:94).

In a later statement in the same article, Green adequately summarizes the controversy over linear versus configural judgment models.

> Whether one accepts the first approximations as good descriptions of reality or as fictions contributed by the method of analysis depends partly on one's purposes. If the goal is prediction in some practical situation, an adequate description will serve. But if the goal is to understand the process, then we must beware of analyses that mask complexities (Kleinmuntz, 1968:98).

The Analysis of Variance Model (ANOVA)

While the preceding discussion has focused primarily on the use of multiple regression techniques, it could just as easily have been formulated in terms of the fixed model analysis of variance (ANOVA). Both systems are simply alternative formulations of a general linear model. The structural elements underlying both the multiple regression and the

ANOVA model are formally equivalent. However, the ANOVA model provides the added capability for making statistical inferences concerning cue interaction (Hoffman, Slovic, and Rover, 1968).

When judgments are analyzed in terms of the ANOVA model, a significant main effect for cue X_1 implies that the decision-maker's responses varied systematically with X_1 as the levels of the other cues were held constant. Similarly, a significant interaction between cues X_1 and X_2 implies that the decision-maker was responding to particular patterns of those cues (i.e., the configured effect of variation of cue X_1 upon judgment differed as a function of the corresponding level taken by cue X_2). It is possible to calculate an index of the importance of individual or configural use of a cue, relative to the importance of other cues. The index, ω^2, described by Hays (1973) provides a rough estimate of the proportion of the total variation in a person's decisions which can be predicted from a knowledge of the particular levels of a given cue or of a configural pattern of cues.

The ANOVA model was also utilized in this study to determine if interactions existed among the cues provided in the research experiment.

Applications of Linear Models for Judgment Modeling

During the past five years, linear judgment models (or policy capturing models) have been used with a great deal of success to analyze complex real-world judgments. Judges in these studies have included business managers, graduate admissions committees, auditors, accountants, loan officers, literary critics, and trout hatchery employees as they attempted to predict business failures and stock performance, select graduate students, plan work force and production schedules, evaluate accounting procedures, evaluate theatrical plays, and recommend trout

streams. Even United States senators have been modeled and their roll-call votes predicted (Slovic, et al., 1976:21). Linear equations have predicted these complex judgments quite accurately.

A research study that parallels the approach taken by this study was conducted at the United States Air Force Academy. The Academy study explored the usefulness of the policy capturing technique in understanding and improving an important appraisal process -- the semi-annual ratings received by the Academy cadets.

The formal evaluation procedure calls for each cadet to be rated on ten performance dimensions: performance of duty, initiative, judgment, human relations, expression, personal moral standards, personal appearance, acceptance of authority, cooperation, and leadership. An overall rating is then calculated by the rater. Each rater must consider the performance dimensions, but there is no scheme prescribed, such as averaging or summing the performance variables.

A multiple linear regression model was used to determine the consistency of raters in applying their rating policies. The findings of the study were as follows: 1) raters' stated policies differed widely from the policies they actually employed as identified through the policy capturing technique; 2) raters were internally consistent in applying their individual actual policies (the policy as captured); 3) both stated and actual policies varied widely between raters; and, 4) overall ratings could be predicted using only two or three of the ten cues employed in the formal rating procedure. Also, the linear model explained as much of the variance in decisions as did a nonlinear model.

The study was replicated with similar results. The question was then raised, If the overall performance rating was predictable on the

basis of only two of the ten cues, might the Academy policy for determining overall ratings be violated significantly? Generally, an equal weighting of all performance variables was assumed by policy makers as a basis for the overall rating.

As a result of these studies, a new rating form was developed for Academy test application. The ten cues were retained. However, the raters only score the cue variables. The overall rating is computed on the basis of uniform weights applied to the cues. The weights are developed by the Academy leaders, with the weights differing for each class, reflecting the training objectives of the respective classes (Taylor and Wilsted, 1976).

Commonalities across the Applications

The results of the Air Force Academy study are consistent with those across a number of studies with varying numbers of cues. Three cues usually account for about 80% of the predictable variance in judges' responses. In many studies, the most important cue usually accounted for 40% of the variance. In addition, judges appear to be unable to state their judgment policy and then apply that policy consistently to their decisions (Slovic and Licktenstein, 1971:680).

Also, as the Academy study pointed out, although individual judges are internally consistent in applying a policy, inter-judge consistency in applying a policy varies widely.

Summary

This chapter reviewed some of the existing literature regarding the modeling of human judgment. Although the process of human judgment is complex, researchers feel that the manner in which humans use available information in making decisions can be modeled.

The overall approach which research has taken in the area of modeling human judgment can be categorized under two frameworks - the Bayesian and the regression. Because this study utilizes the regression framework, the discussion centered on the use of regression techniques. A rather detailed description of the multiple linear regression model was presented to familiarize the reader with the methodology used in this study.

To dispell the conclusion that linear models are the panacea for judgment modeling, the arguments for using nonlinear or configural models were discussed. There are of course judgment processes that do involve configural use of the available cues. In such cases the linear model (ANOVA) is useful for identifying any interaction in the utilization of the cues by a judge.

Finally, applications of linear models attest to the utility of these models in accurately predicting judgment processes across a wide spectrum of real-world situations.

III Research Methodology

Policy capturing, as it is considered in this study, is a deductive analytical technique. It entails the investigative strategy utilized by scientists testing hypotheses of general scientific interest. As a consequence, the methodology follows a rather structured framework for developing the experimental design and analyzing the collected data from the experiment.

As a general overview, this study utilized a decision-making exercise which was given to junior and senior ranking Air Force officers. Each officer was asked to evaluate, on a scale of one to seven, the promotability of 32 hypothetical Air Force captains. Policy capturing techniques were used in the design of the exercise and the analysis of the resulting data. This chapter describes how the experiment was designed and reviews the various analyses which were used to process the data.

The Specific Decision to be Modeled

In order to investigate the whole person concept of Air Force promotions, it was felt that some type of promotion decision should be modeled which would provide a measure of the relative importance individuals place on specific promotion factors. For the most part, the interests of the writer (an Air Force captain) determined the final decision to be modeled. As a result, the specific decision centered on the promotability of Air Force captains.

As with any decision, the evaluation of individual promotability requires that the decision-maker (judge) look at specific information or cues about the individual and then render a decision. The cues (or factors of promotion) which were to be included in the experimental design were, therefore, very important.

Identification of Factors for Decision Criteria

As it was pointed out in Chapter I, the Air Force does not specifically state a policy regarding what factors are important for promotion. As a result, the search for such factors was relatively subjective in nature.

There are probably many factors which influence a promotion decision. However, because of the specific experimental design of the decision-making exercise, which is discussed in a later section, five factors of promotability were chosen. The factors chosen were: 1) Officer Effectiveness Report ratings, 2) formal education, 3) professional military education, 4) assignment history, and 5) aeronautical rating. By utilizing only five factors, the decisions were kept relatively simple without detracting from the realism of the decision.

Although these factors were selected subjectively, there is support for such factors in the literature. The October 18, 1976 major selection board selected 942 out of 959 (98%) new eligibles with a "1" rating on their Officer Effectiveness Reports (OER). Of the 960 officers with a "2" rating, 806 were selected (84%). And 252 of 855 (29%) with a "3" rating were promoted to major (Ewing, 1977:3). Obviously, OER ratings are an important promotion factor.

An Armed Forces Journal International article indicates that assignment level has a lot to do with promotions, particularly for promotion to lieutenant colonel. In the 1973 selection process, most selectees were assigned to high Air Force staff jobs or at major Air Force Commands ("Short Sheeted Again," 1973:14).

In addition, General John C. Meyer, Commander and Chief, Strategic Air Command, stated that promotion boards stress the "whole man concept."

He lists three key factors: 1) performance, 2) education, and 3) job diversification. OERs provide the primary measure of performance. Education includes professional military education (PME). General Meyer further states, "...there is a direct correlation between formal educational level and promotion to major, lieutenant colonel, and colonel." In addition he states, "Completing rungs on the PME ladder can be a key separator at selection time" (Meyer, 1973:1-2).

The aeronautical rating factor was probably determined most subjectively. Initially, the researcher determined six factors; the five finally selected plus military decorations. However, the use of six factors would have complicated the decision exercise because of the strict requirements of a full-factorial experimental design (the full-factorial design is discussed in detail in a following section).

In an attempt to limit the number of factors to five, an informal survey of 46 officers attending the Air Force Institute of Technology was conducted. These officers were given a piece of paper with the above mentioned five factors plus military decorations as a sixth factor. They were asked to circle the factor which they felt was the least important for promotion to major. Of the 46 officers surveyed, 44 perceived that military decorations were least important of the six listed factors in a promotability decision. Consequently, aeronautical rating was included in the exercise.

Design of the Decision-Making Exercise

The decision-making exercise developed for this study can be technically described as a full-factorial experimental design. Factorial designs require that strict choices be made for the combinations of cue values for each case. To construct a factorial experiment, the following procedure is used.

First, several values or levels for each cue (promotion factors) are chosen. In this case two levels were chosen for each promotion cue or factor. They were as follows:

OER Ratings - In a series of three OERs an individual either received ratings of 2, 2, 1 or ratings of 2, 1, 1.

Formal Education - An individual had completed either a masters degree or a bachelors degree.

PME History - An individual had either completed Squadron Officers' School or had no PME on record.

Assignment History - An individual had completed a headquarters assignment or else assignments had been entirely at the base level.

Aeronautical Rating - An individual was either a pilot or a nonpilot.

Concerning the OER ratings factor, it is important to understand that the depicted ratings conform to the new OER policy announced by Headquarters Air Force in August 1977. This means, for the OER ratings used in this study, 22% of the officers could receive an overall rating of "1" while 78% of the officers could receive an overall rating of "2" or lower.

The reader should note that the two possible combinations of OER ratings (levels) are not significantly different. In fact the two levels only differ with respect to one OER. This structure was deliberately utilized in order to elicit the whole person concept of evaluation from the experimental subjects. It was felt that cases involving similar performance ratings among eligible officers justify the use of the whole person concept of evaluation.

The second step in a factorial design is to construct cases by combining each level of each cue with every chosen level for the remaining

cues. This process then gives one case for every possible combination of levels for all cues. This results in fixing the total number of cases to be considered. If N is the total number of cases to be considered, N is computed by

$$N = N_1 \times N_2 \times N_3 \times \ldots N_k \qquad (2)$$

where N_i is the total number of levels associated with each of k cues.

In this case the exercise task has five cues, each having two levels. Therefore

$$N = N_1 \times N_2 \times N_3 \times N_4 \times N_5$$

$$N = 2 \times 2 \times 2 \times 2 \times 2$$

or

$$N = 32 \text{ cases}$$

Factorially arranged cues allow an analyst to employ many useful statistical tests. The important feature of factorial experiments is that the cues are orthogonal (independent). This leads to an important statistical property: The cues and all cue interactions are uncorrelated. This property allows an analyst to draw cause-and-effect relationships from the results of an experiment. For example, if a certain cue is found to be significant in predicting the response, the hypothesis that the cue is utilized by the decision-maker is confirmed.

One drawback to the use of a factorial design is the large number of cases required. Obviously, for the five cues which were chosen there are actually more than just two levels. However, for each level that is added, a considerable number of cases are generated which in

turn complicates the decision task. Therefore, as stated previously, this decision-making exercise was limited to five factors, each with two levels in order to simplify the task.

The decision exercise used for this study is presented in Appendix A. The final instrument was the result of several modifications and interactions. The initial instrument was administered to 12 Air Force Institute of Technology students as a test. Several suggestions from that administration were incorporated. In addition, further modifications resulted from suggestions received from higher headquarters' personnel who reviewed the exercise.

The 32 cases were randomly arranged in the exercise booklet by assigning an arbitrarily selected random number from a random number table to each case. The random numbers and the cases associated with each number were then listed sequentially. It should be noted that within each case, the presentation of the five cues was also randomized. This was done to avoid the possibility that an individual might unconsciously assign the most weight to the factor presented first for each case.

<u>Relating the Exercise to the Linear Regression Model</u>

Recalling the general form of the multivariate linear regression model discussed in Chapter II, the appropriate model for this study is

$$Y = b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5 + \varepsilon \qquad (3)$$

where X_1 through X_5 are the predictor variables translated from the selected promotion factors as follows:

X_1 = PME History

X_2 = Assignment History

X_3 = Aeronautical Rating

X_4 = OER Ratings

X_5 = Formal Education

ε = Error Term

Each decision case contains information concerning the level of each of the predictor variables as shown in Fig. 3.1.

CAPTAIN #1

DECISION CRITERIA	INDIVIDUAL PERFORMANCE
PME HISTORY	Squadron Officers' School
ASSIGNMENT HISTORY	Entirely at base level
AERONAUTICAL RATING	Pilot
OER RATINGS	2, 1, 1
FORMAL EDUCATION	Masters degree

DECISION #1

1 . . . 2 . . . 3 . . . 4 . . . 5 . . . 6 . . . 7

Low Promotability — Moderate Promotability — High Promotability — Very High Promotability

Fig. 3.1. Sample Decision Case

As discussed earlier, the use of two levels for each predictor variable gives the model special status since the 32 decision tasks can each be represented by a mathematically orthogonal vector. Each of the 32 evaluations made by the experimental judge is strictly characterized by the numerical value (1, 2, ..., 7) of the evaluation and one of 32 orthogonal predictor vectors, each of which represents a unique performance

state. In this study, the predictor variables were assigned a value of either "0" or "1," depending on the specific level presented for each case as depicted in Table 3.1.

Table 3.1
Numerical Values Assigned to Predictor Variables

Predictor Variable	Level	Value Assigned Predictor Variable
PME History (X_1)	SOS completed	1
	No PME	0
Assignment History (X_2)	Headquarters tour completed	1
	Entirely at base level	0
Aeronautical Rating (X_3)	Pilot	1
	Nonpilot	0
OER Ratings (X_4)	Ratings of 2, 1, 1	1
	Ratings of 2, 2, 1	0
Formal Education (X_5)	Masters degree	1
	Bachelors degree	0

In essence, the evaluation number (1, 2, 3, ..., 7) corresponds to the criterion variable Y in the model. The predictor variables are assigned a value of 1 or 0 depending on the level of the variable given for each case. As a result, for each exercise respondent a set of 32 linear equations is developed. These equations are solved simultaneously to yield values for the standardized regression coefficients (b_i).

Once the value for the regression coefficients (b_i) have been calculated, a measure of the relative importance placed on each factor (predictor variable) can be computed from a formula derived by Hoffman (1960).

The formula is

$$W_i = \frac{b_i^2}{R^2} \qquad (4)$$

where

W_i = relative weight associated with the ith predictor variable (promotion factor)

b_i = standardized regression coefficient for the ith predictor variable (promotion factor)

R^2 = squared multiple correlation coefficient for the model

The above formula applies to experimental designs that provide for orthogonal vectors such as the one used in this study.

Therefore, by administering the decision-making exercise to a selected group of individuals, it is possible to extract, among other things, the relative importance the group or the individual places on each promotion factor.

Collection of Data

The source of data for this study came from Air Force professional military education (PME) schools. These schools were the Squadron Officers' School (SOS), Air Command and Staff College (ACSC), and Air War College (AWC). It was felt that these three schools would provide data consistent with the objectives of the study. Sample data from SOS would be representative of junior officer views, data from ACSC would provide the viewpoint of recently promoted majors, and AWC would provide data relevant to the view of senior-ranking Air Force officers, some of whom may well sit on future promotion boards.

The decision exercises were personally delivered to each of the three PME schools. A designated office within each school became the

focal point for the distribution and collection of the exercises. The designated office for AWC was the Requirements Office. In the case of ACSC and SOS, the designated offices were the respective Evaluation Branches. Personnel in each of the designated offices assisted in randomly selecting officers, distributing the exercises to these selected officers, and collecting the completed instruments. The completed exercises which were not collected within a three-day period were later returned by mail in bulk.

Only active duty Air Force officers were identified to participate in the exercise. Air Force Reserve, Air National Guard, and officers of other services were eliminated from the sampled population. This left a total population of 677 SOS officers, 420 ACSC officers, and 196 AWC officers from which 160 officers were randomly selected from each school. A total of 318 useable exercises were returned for a response rate of 66.3%.

Coding of Collected Data

Each completed exercise was marked with an alphanumeric code which identified the PME school and that specific exercise. Data from the completed exercise were coded onto a standard IBM data card. Fig. 3.2 shows the card format and variable names used for the data base. The name and address of the officer was recorded on a separate data card only if the individual requested an analysis of performance in the exercise. This option was offered as an incentive to increase the return rate. An example of the feedback provided to those requesting it is presented in Appendix C.

Column Number	Entry
1-3	School identification
4-6	Exercise number
7	Rank
8	Aeronautical rating
9	Present assignment
10	Highest assignment
11	Level of education
12	Number supervised
13-15	Last three controlled OERs
16-47	Individual decisions
48-58	Subjective weights

Fig. 3.2. Key to Coding of Exercise Data on Data Cards

Restructuring the Data for Regression Analysis

Each coded data card contained the numerical values (1, 2, 3, ..., 7) for every promotability decision made by the sampled officers. These data cards did not, however, contain the predictor vector associated with each of the 32 decisions. In addition, the 32 decisions were coded on the cards in a horizontal format, but the computer regression algorithm required that the regression variables be read in line by line. Therefore, it was necessary to restructure the data for regression analysis.

Previous research efforts in policy capturing required similar analysis techniques. As a consequence a FORTRAN program was available, which had been written by Major C. W. McNichols, a professor in the Air Force Institute of Technology Systems Management Department. This program was modified to meet the requirements of this study.

The program reads the numerical values for the 32 decisions coded on each card and creates an $n \times 6$ matrix for storage of each decision and its respective predictor vector. The value n is 32 times the number of officers who responded to the exercise. The program adds the appropriate

predictor vector to each evaluation and then stores the entire matrix on a disk file. In addition, the program files the individual demographic data with each case.

Computational Aids Used in the Analysis of Data

Most analyses for this study were performed with the aid of the Aeronautical Systems Division (ASD) CDC 6600 computer at Wright-Patterson Air Force Base, Ohio. A majority of the computer analyses employed the Statistical Package for the Social Sciences, commonly called SPSS (Nie, et al., 1975). SPSS is a library of statistical analyses programs which are resident in the computer. The mechanics of utilizing SPSS have been designed so that even an individual with very little programming experience can quickly learn to use the system.

In addition to SPSS, another computer program, which was also written by Major McNichols, was modified for use in this study. The program calculates the relative weights and the multiple correlation coefficient for every respondent to the exercise and files that information on a disk file for later analyses.

Additional computations were made "manually" on a Texas Instrument SR51-II calculator. These computations included the calculation of F_o values for the tests of significance for differences between the group regression models. These procedures will be explained further in following sections of this chapter.

Analyses Performed on the Data

The following sections describe the analyses which were performed on the exercise data. In general, the promotability decisions from the exercise were modeled through the use of regression analysis and analysis of

variance (ANOVA). The additional analyses which are discussed were used to test specific hypotheses concerning the three separate groups of officers who responded to the decision exercise.

Descriptive Statistics. The first analysis utilized the SPSS subprogram frequencies: one-way frequency distributions with descriptive statistics (Nie, *et al.*, 1975:194). This analysis provided an easy means for inspecting the whole data base. The output of the program gave the number of respondents for each question. As an option to the program, it was possible to have histograms printed out which graphically depicted the distribution of the data for each question. This analysis was a valuable tool for detecting errors in the data.

Group Regression Analyses. The SPSS regression algorithm is extremely flexible. It provides several types of regression, tests for significance, and various statistics. The data for this study was subjected to stepwise regression in every case. In every regression the promotability decision was regressed on X_1 through X_5 (the promotion factors).

Group regressions were run for the following groups:

Run 1: All groups combined
Run 2: SOS and ACSC combined
Run 3: SOS and AWC combined
Run 4: ACSC and AWC combined
Run 5: SOS
Run 6: ACSC
Run 7: AWC

The output of the SPSS regression program provided the following information: 1) the group R^2 for each run, 2) the standardized regression coefficients (beta weights), 3) the group residual sum of squares for later use in the F-test of significance, and 4) the F-test level of significance for each promotion factor.

Individual Regression Analysis. Regression analysis for each individual respondent was accomplished by utilizing the previously mentioned FORTRAN program. The output of this program provided the individual relative weights placed on each factor and the individual R^2 value. This data base was then used extensively for additional SPSS analyses.

Analysis of Variance (ANOVA). Analysis of variance was utilized to detect any significant interactions among the promotion factors. The ANOVA methodology was possible because of the full-factorial design of the exercise.

Once again SPSS provided the program for accomplishing the ANOVA. The output of the program depicts the main effects and two-way, three-way, four-way, and five-way interactions. It also provides the level of significance for each of these interactions. For this analysis, an n-way ANOVA was accomplished with the promotability decision as the dependent variable and the five promotion factors as the independent variables.

F-Test of Significance. In order to determine if there was a statistical difference in the way the different groups of officers weighted the promotion factors, the F-test as described by G. C. Chow (1960:599) was utilized. The null hypothesis tested in each comparison is that there is no significant difference in the regression coefficients of the models being compared. The alternate hypothesis is that there is at least one model among those being compared whose regression coefficients are significantly different from the others for at least one pair of regression coefficients. Rejection of the null hypothesis, for purposes of this study, is equivalent to saying that for the groups being compared there are significant differences in the weights at least one of the groups places on the five promotion factors. It should be noted that this

analysis does not identify the specific factors that differ, however. The level of significance used for all F-tests used in this study was 0.05.

The F_o value (the observed F value) was computed "manually" using data provided by the SPSS group regressions. The method of calculating the F_o value is described in Appendix B for the interested reader.

Student t-Test of Significance. Student's t-test, as described by Freund (1971:317-318), was used to further analyze the individual regressions. This analysis was used to test the null hypothesis that the mean promotion factor relative weights (computed from the individual regressions) were the same for the three groups of officers. The alternate hypothesis was that these means were not the same for one or more of the groups. In order words, the test shows which promotion factors are weighted significantly different among the three groups of officers. The level of significance used for this test was 0.05.

Paired Samples t-Test. The final task in the decision exercise asked each individual to indicate the relative importance placed upon the five promotion factors by distributing 100 points to the factors. This, in effect, asked the individual to state his/her decision policy. From henceforth, these individual weights will be referred to as the subjective weights. The actual relative weights determined from regression analysis will be termed the objective weights.

The paired samples t-test was utilized to determine if, as a whole, the groups were able to apply their stated decision policies to the decisions in the exercise. In essence, the paired t-test compares the mean objective weight for a specific factor to the corresponding subjective weight. To do this, the paired difference variable $D = S-O$ is formed, where S is the subjective weight and O is the objective weight.

D is normally distributed with mean δ. The null hypothesis is then $H_0: \delta = 0$, and the alternate hypothesis is $H_1: \delta \neq 0$. The level of significance for this test was 0.05.

Both of the t-tests described above are subprograms contained in SPSS.

R^2 as a Measure of Consistency

The fact that this study incorporated a full-factorial experimental design means that the models involved have orthogonal predictor variables. This provides a unique interpretation for R^2, which is termed the squared multiple correlation coefficient for the regression equation. If a regression equation is generated for each individual judge making the same decision, using the same information cues, such an equation expresses the policy of that judge (Christal, 1968:26).

In the case where the predictor variables are uncorrelated, which is the case in this study, the value of R^2 can actually be viewed as a measure of the consistency of a judge in applying a decision policy to the 32 promotability decisions.

It should be noted that to this point the interpretation of R^2 has been related only to the individual regression model. When this notion is extended to models representing larger groups (such as the group regression model), the same interpretation of R^2 applies. However, the consistency measure associated with R^2 is not quite so clear for two reasons: 1) a greater number of random errors is introduced in larger groups and 2) each judge may be very consistent in applying an individual policy, but the composite group may not reflect the same degree of consistency as the individual judges (O'Berry, 1977:40).

Suppose an individual regression model produces large R^2 values for each judge in a large group of judges. However, when this same group of judges is modeled by a group regression, the overall R^2 value for the model is much lower than the mean individual regression R^2 for the group. In this case, one can conclude that the individual judges are consistent in applying their decision policies. However, the group regression model indicates that the decision-making policies within the group as a whole are not homogeneous; or, there is a great deal of variety in the way individuals in the group make their decisions.

Summary

This chapter covered the research methodology incorporated in this study. The overall approach of the research was to develop a decision-making exercise that required an individual to evaluate the promotability of 32 hypothetical Air Force captains. Policy capturing techniques were then applied to analyze the data.

The basic design of the decision exercise followed a full-factorial experimental design. This, in turn, insured that the available cues for the decision were orthogonal. In addition, because the exercise was a full-factorial design, only five cues with two levels per cue were included in the exercise.

The cues were limited in an attempt to keep the exercise as simple as possible without sacrificing the realism of the decisions. The final five promotion factors which were included in the exercise as decision criteria were: 1) professional military education, 2) assignment history, 3) aeronautical rating, 4) Officer Effectiveness Report ratings, and 5) formal education.

The decision exercise was administered to three separate groups of active duty Air Force officers attending the Air War College, Air Command and Staff College, and Squadron Officers' School, all located at Maxwell Air Force Base, Alabama. Through the use of regression analyses, it was possible to determine the relative weight each group placed on the five promotion factors.

Additional analyses were discussed which basically allowed for testing hypotheses concerning the differences or similarities between the groups of officers and their perceptions of what is important for promotion.

IV Research Results

This chapter summarizes the results of the various analyses which were discussed in Chapter III. In the interest of the reader, it may be beneficial to review the overall analysis approach since it is often easy to lose sight of the objective when beseiged with descriptions of statistical techniques.

The fundamental objective in the analysis of the data was to derive quantitative values that reflect the relative weights Air Force officers place on the five factors of promotion used in this study. With these factor weights, it is then possible to utilize statistical inference for testing various hypotheses concerning the differences and/or similarities between the three groups of officers who participated in the exercise.

The basic model used to derive the relative weights which were placed on the promotion factors was the linear regression model. In essence, this model was used in two ways: First, to model the promotability decisions on the basis of the three officer groups and, second, to model the decisions of each individual officer. For the most part, the additional analyses discussed in Chapter III statistically tested for differences among the three groups.

With this reaffirmation of the overall analysis objective, the chapter begins with an overview of the research data base.

The Decision Exercise Response Rates

Table 4.1 summarizes the overall response to the decision-making exercises which were distributed to the three PME schools. As depicted, 480 total exercises were evenly distributed between the three schools and 318 officers chose to participate. This resulted in an overall response rate of 66.3%.

Table 4.1
Decision Exercise Response Rates

PME School	Number of Exercises Distributed	Number of Useable Returns	Response Rate
SOS	160	99	61.9%
ACSC	160	134	83.8%
AWC	160	85	53.1%
Total	480	318	66.3%

In an attempt to increase the response rate, the exercise participants were offered the option of receiving a summary comparison of their results with those of their peers. It is interesting to note that 81.5% of the respondents desired feedback. This fact, in addition to numerous written comments returned with the exercises, indicated that there was keen interest in this research.

The Distribution of Officer Rank by PME Schools

All analyses performed on group data were accomplished by classifying three groups of officers. These three groups corresponded to the three PME schools. Table 4.2 depicts the rank structure within each of the schools based upon the officers who returned the exercises.

It is important to realize that the 45 captains who responded from the Air Command and Staff College were actually major selectees. This was the primary reason for performing the analysis by PME schools as opposed to grouping by officer rank.

Table 4.2
Distribution of Officer Ranks Within
PME Schools for Exercise Respondents

Rank	SOS	ACSC	AWC
First Lieutenant	23	0	0
Captain	74	45*	0
Major	0	86	0
Lieutenant Colonel	0	0	79
Colonel	0	0	6
Unidentified	2	3	0
Total	99	134	85

* Major selectees

Table 4.3 provides a breakdown of the rated and nonrated officers who responded to the decision exercise.

Table 4.3
Distribution of Rated and Nonrated
Officers for Exercise Respondents

PME School	Number of Pilots	Number of Navigators	Number of Nonrated Officers	Missing Cases
SOS	38	20	39	2
ACSC	43	11	79	1
AWC	50	14	20	1
Total	131	45	138	4

Results from the Group Regression Model

As explained previously, the group regression model takes all the decisions of officers from a specific group (school) along with the associated predictor vectors and calculates the standardized regression

coefficients (beta weights). This in turn yields a linear equation which is, in fact, a model of the way that a specific group of officers made their decisions. In addition, the regression algorithm computes the squared multiple correlation coefficient (R^2). The numerical value of R^2 is the percentage of the total variations in the decision making of the group which is explained by that specific linear equation.

Once the beta weights have been calculated, it is then possible to compute the relative weights placed on each of the five factors. This computation can be accomplished from either the group regression model or the individual regression model. In order to avoid confusion, the researcher chose to present only one set of values for the relative weights of the promotion factors. These values were computed from the individual regression model, as will be explained later in the discussion of the individual regression analysis results.

Table 4.4 summarizes the results of the group regression analysis for each of the three officer groups.

Table 4.4
Summary of Group Regression Analysis

Officer Group	All Five Promotion Factors Significant at .001 Level?	Group R^2
SOS	Yes	.439
ACSC	Yes	.382
AWC	Yes	.397

The flexibility of the SPSS regression algorithm provides for an F-test of significance for the individual regression coefficients (unstandardized). The second column of Table 4.4 provides the results

of the F-tests applied to the promotion factor coefficients. In essence, the test showed that each of the five promotion factors were statistically significant in the model for each of the three officer groups. In more basic terms, each of the groups used all five promotion factors in arriving at a promotability decision.

It is very interesting to note the relatively low values of R^2 for the three groups. As will be discussed later, these low values for the group regression R^2, when compared to the mean individual regression R^2 values, demonstrate that there is a wide variety in the way individuals arrive at their decisions regarding promotion.

Utilizing additional numerical data obtained from the SPSS regression program, it was possible to apply an F-test of significance to determine if the three group regression models were statistically different.

F-Test of Significance on Group Regression Models

The F-test as described by Chow (1960) is a useful statistical test for comparing two or more regression models. In this case the F-test was used to compare the three regression equations (models) associated with each of the three PME schools. In effect, this test indicates whether any one pair of regression coefficients is statistically different among the three models. It is important to note that this test only indicates that the overall models are statistically different for at least one pair of beta weights. It does not specify which coefficients are different.

The F-test of significance results are tabulated in Table 4.5. First of all, all three models were compared. The results indicate that the null hypothesis should be rejected; or at least one of the models is statistically different from the other two models. The level of significance for all F-tests is .05.

Table 4.5
F-Test Results

Comparison Groups	N	F_o	$F_{.05}$	Reject H_o?
AWC/ACSC/SOS	10,176	26.46	1.75	Yes
AWC/ACSC	7,008	3.62	2.10	Yes
AWC/SOS	5,888	41.95	2.10	Yes
SOS/ACSC	7,456	23.83	2.10	Yes

The next series of F-tests compares the models in pairs. In each case the null hypothesis was rejected. So what does this mean to the analyst?

The results of these series of F-tests indicate that each of the officer groups statistically differ on <u>at least</u> one beta weight from the other two PME schools. For example, in comparing SOS and AWC perhaps the AWC group of officers placed significantly more weight on OERs than did the SOS officers. Once again, however, the F-test <u>does not</u> indicate the specific regression coefficients that differ (such as OERs). Additional analysis is required to identify the specific differences.

Thus far one can now infer that, yes, the three groups of officers differ in some respect regarding the importance placed on the five factors of promotion. In order to further identify the specific differences, the individual regression model must be utilized in conjunction with the Student t-test of significance.

Results from the Individual Regression Model

As stated previously, the individual regression model provides a linear equation that models the decisions for each individual officer.

From the beta weights computed by the regression algorithm the relative weights for each of the five promotion factors were calculated for each individual officer. The relative weights were computed by utilizing the Hoffman relative weight measure discussed earlier.

Once the individual relative weights were determined, it was possible to compute the mean relative weights placed on the five promotion factors for a specific group of officers. These mean relative weights, classified according to PME school, are depicted in Table 4.6.

Table 4.6
Results from Individual Regression Model
(Mean Values)

Officer Group	Promotion Factor	Mean Relative Weight	Group Mean R^2
SOS	OER	.200	.762
	Formal Education	.236	
	PME History	.349	
	Assignment History	.128	
	Aeronautical Rating	.087	
ACSC	OER	.325	.776
	Formal Education	.194	
	PME History	.292	
	Assignment History	.135	
	Aeronautical Rating	.054	
AWC	OER	.414	.791
	Formal Education	.127	
	PME History	.226	
	Assignment History	.146	
	Aeronautical Rating	.087	
All Groups	OER	.310	.777
	Formal Education	.189	
	PME History	.292	
	Assignment History	.136	
	Aeronautical Rating	.073	

By inspecting the tabulated relative weights, one can see that the three groups placed different weights on the five promotion factors. In

order to clarify the comparison of the three groups for the reader, the graphical depiction of the mean relative weights is presented in Fig. 4.1. The reader should be cautioned that the connecting lines in Fig. 4.1 are not meant to imply a continuous function between the five promotion factors. The lines merely aid in associating the five factors with a specific PME school.

At this point, it is appropriate to point out that the large relative weights placed on the PME factor by all three officer groups is not too surprising, considering the fact that all the experimental subjects were enrolled in PME schools.

Although it is apparent that the three officer groups placed different weights on the five factors, it has not been determined that the groups are statistically different. The Student t-test of significance was used to make this determination.

Results of the Student t-Test of Significance

The Student t-test was used to determine whether or not the relative weights placed on the five promotion factors were significantly different between groups. "Significant" here does not imply "important" or "consequence": it is used here to mean "indicative of" or "signifying" a true difference between two groups of officers (Nie, et al., 1975:267).

Tables 4.7 through 4.9 summarize the findings associated with the t-test of significance performed on the relative weights. The reader should note that each of the paired group comparisons (e.g., SOS compared to ACSC) shows that the respective officer groups differ consistently on the relative weights placed on three promotion factors: 1) PME, 2) OER ratings, and 3) formal education. This result will be discussed further in the next chapter.

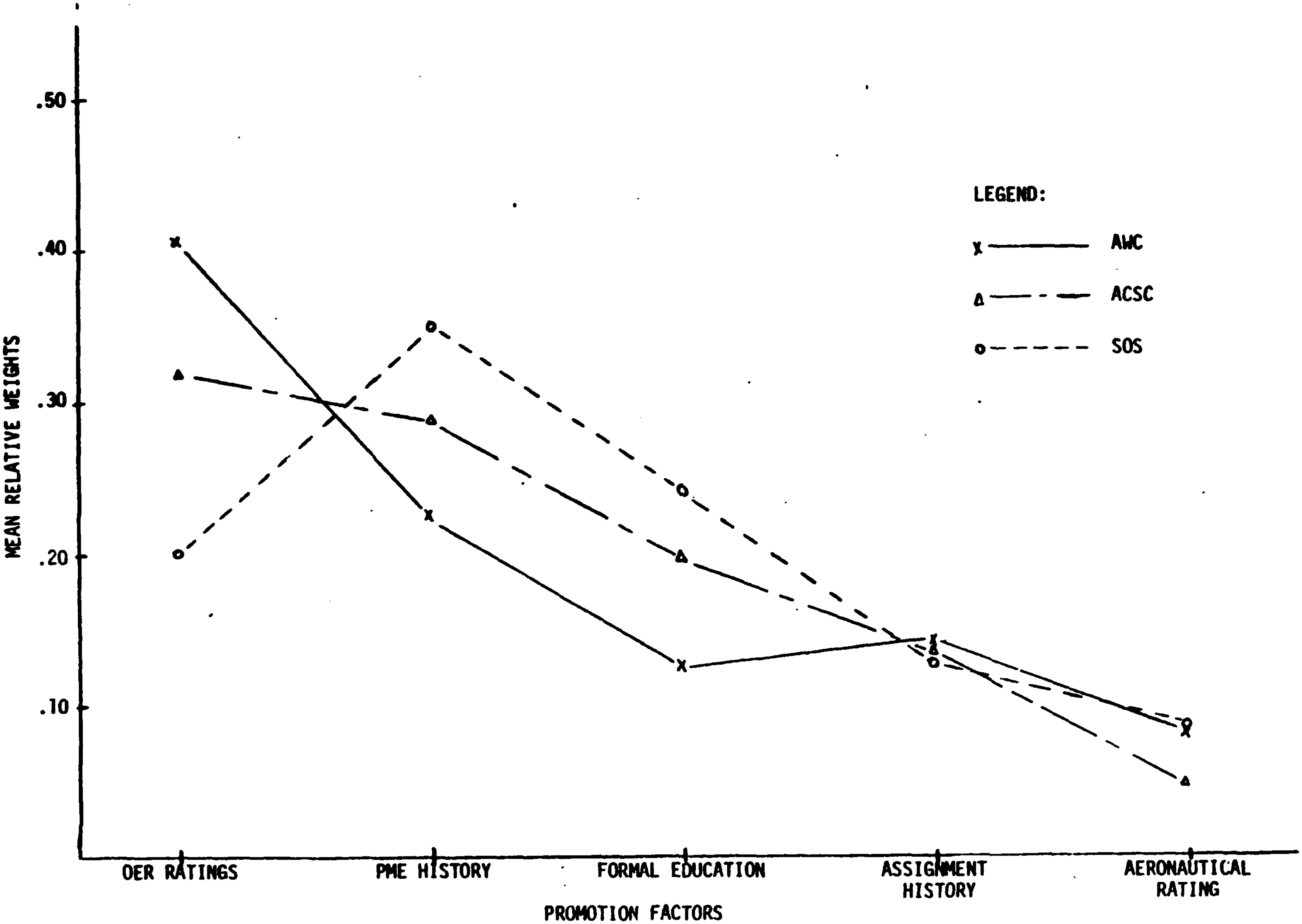

Fig. 4.1. Mean Relative Weights Placed on Factors Based on Individual Regression Model

Table 4.7
t-Test of Significance Comparing
Relative Weights Between SOS and ACSC Officers
(Level of Significance is .05)

Promotion Factor	Officer Group	Mean Relative Weight	Statistically Different?
PME	SOS	.349	Yes*
	ACSC	.292	
Assignment History	SOS	.128	No
	ACSC	.135	
Aeronautical Rating	SOS	.087	Yes*
	ACSC	.054	
OER Ratings	SOS	.200	Yes**
	ACSC	.325	
Formal Education	SOS	.236	Yes**
	ACSC	.194	

* p < .05
** p < .001

Table 4.8
t-Test of Significance Comparing
Relative Weights Between SOS and AWC Officers
(Level of Significance is .05)

Promotion Factor	Officer Group	Mean Relative Weight	Statistically Different?
PME	SOS	.349	Yes*
	AWC	.226	
Assignment History	SOS	.128	No
	AWC	.146	
Aeronautical Rating	SOS	.087	No
	AWC	.087	
OER Ratings	SOS	.200	Yes*
	AWC	.414	
Formal Education	SOS	.236	Yes*
	AWC	.127	

* p < .001

Table 4.9
t-Test of Significance Comparing
Relative Weights Between ACSC and AWC Officers
(Level of Significance is .05)

Promotion Factor	Officer Group	Mean Relative Weight	Statistically Different?
PME	ACSC	.292	Yes*
	AWC	.226	
Assignment History	ACSC	.135	No
	AWC	.146	
Aeronautical Rating	ACSC	.054	No[1]
	AWC	.087	
OER Ratings	ACSC	.325	Yes*
	AWC	.414	
Formal Education	ACSC	.194	Yes**
	AWC	.127	

* $p < .05$
** $p < .001$
[1]Separate variance estimate used instead of pooled variance. Level of significance was .055.

The ANOVA Model Findings

Tables 4.10 through 4.13 tabulate the results of applying an n-way analysis of variance to each of the three officer groups. Table 4.10 summarizes the ANOVA findings for SOS officers. The column labeled "Percent of Explained Variation" shows that the five promotion factors, when considered as the main effects, account for 98.4% of the explained variations in the decisions of SOS officers. Similarly, significant (at the .05 level) two-way interactions account for only 0.7% of the explained variance. In other words, the five promotion factors, when combined in a linear fashion, explain practically all of the explained variance in the model.

Similar results are shown in the remaining tables. Once again, the findings of the ANOVA analysis confirm that all five factors were used by the officers in making their promotability decisions. This is reflected by the column showing the F-test level of significance.

Table 4.10
Summary of Factor Interactions
for SOS Based on ANOVA Analysis
(Statistical Significance at .05 Level)

Source of Variation	F-Test of Significance	Percent of Explained Variation
Main Effects:		
PME History	.001	40.6
Assignment History	.001	9.9
Aeronautical Rating	.001	5.8
OER Ratings	.001	17.6
Formal Education	.001	24.5
Total:		98.4
Two-Way Interactions:		
PME History with OER Ratings	.005	0.3
PME History with Formal Education	.001	0.4
Others		0.1
Total:		0.8

Table 4.11
Summary of Factor Interactions
for ACSC Based on ANOVA Analysis
(Statistical Significance at .05 Level)

Source of Variation	F-Test of Significance	Percent of Explained Variation
Main Effects:		
PME Ratings	.001	30.2
Assignment History	.001	12.8
Aeronautical Rating	.001	2.4
OER Ratings	.001	32.6
Formal Education	.001	20.3
Total:		98.3
Two-Way Interactions:		
PME History with Formal Education	.001	0.2
PME History with Assignment History	.033	0.2
PME History with OER Ratings	.013	0.5
Others		0.2
Total:		1.1

Table 4.12
Summary of Factor Interactions
for AWC Based on ANOVA Analysis
(Statistical Significance at .05 Level)

Source of Variation	F-Test of Significance	Percent of Explained Variation
Main Effects:		
PME History	.001	21.8
Assignment History	.001	14.4
Aeronautical Rating	.001	4.7
OER Ratings	.001	45.8
Formal Education	.001	12.4
Total:		99.1
Two-Way Interactions:		
None at .05 level of significance		

Table 4.13 Summary of Factor Interactions for All Groups Based on ANOVA Analysis (Statistical Significance at .05 Level)		
Source of Variation	F-Test of Significance	Percent of Explained Variation
Main Effects:		
PME History	.001	31.5
Assignment History	.001	12.6
Aeronautical Rating	.001	4.1
OER Ratings	.001	30.9
Formal Education	.001	19.7
Total:		98.8
Two-Way Interactions:		
PME History with OER Ratings	.001	0.2
PME History with Formal Education	.001	0.3
Assignment History with OER Ratings	.031	0.1
Others	.031	0.1
Total:		0.7

Comparison of Subjective and Objective Relative Weights

As discussed previously, the final task in the decision exercise requested that the officers document their perceived decision policies by distributing 100 points among the five promotion factors. A paired sample t-test was then used to compare the computed relative weights (objective) with the perceived relative weights (subjective). Again, this specific t-test determines whether there is a significant difference between the subjective and objective relative weights.

Tables 4.14 through 4.16 depict the results of the paired sample t-test by PME schools. Table 4.14 shows that the subjective versus objective mean relative weights for SOS officers were statistically different for four out of the five factors. In order words, the weight these officers thought they put on four of the factors was significantly different from the weight they actually placed on those four factors.

In the case of ACSC officers, Table 4.15 shows that these officers also had significantly different objective and subjective relative weights for four out of the five factors.

For AWC officers, Table 4.16 depicts that these officers had significantly different objective and subjective relative weights for only one factor. Further discussion concerning these findings is deferred until the next chapter.

Table 4.14
Comparison of Subjective and Objective Mean Relative Weights for SOS Officers
(Paired t-Test Level of Significance .05)

Promotion	Type of Mean Relative Weight	Mean Relative Weight	Statistically Different?
PME	Objective Subjective	.349 .223	Yes*
Assignment History	Objective Subjective	.128 .169	Yes*
Aeronautical Rating	Objective Subjective	.087 .101	No
OER Ratings	Objective Subjective	.200 .306	Yes*
Formal Education	Objective Subjective	.236 .201	Yes**

* $p < .001$
** $p < .05$

Table 4.15
Comparison of Subjective and Objective
Mean Relative Weights for ACSC Officers
(Paired t-Test Level of Significance .05)

Promotion Factor	Type of Mean Relative Weight	Mean Relative Weight	Statistically Different
PME	Objective Subjective	.282 .188	Yes*
Assignment History	Objective Subjective	.135 .172	Yes*
Aeronautical Rating	Objective Subjective	.054 .072	Yes**
OER Ratings	Objective Subjective	.325 .386	Yes*
Formal Education	Objective Subjective	.194 .182	No

* $p < .001$
** $p < .05$

Table 4.16
Comparison of Subjective and Objective
Mean Relative Weights for AWC Officers
(Paired t-Test Level of Significance .05)

Promotion Factor	Type of Mean Relative Weight	Mean Relative Weight	Statistically Different?
PME	Objective Subjective	.226 .151	Yes*
Assignment History	Objective Subjective	.146 .164	No
Aeronautical Rating	Objective Subjective	.087 .088	No
OER Ratings	Objective Subjective	.414 .431	No
Formal Education	Objective Subjective	.127 .145	No

* $p < .001$

Additional Findings

Through the use of the flexible data selection options available with SPSS, it was possible to determine the number of officers who weighted each factor the most important. This information is summarized in Table 4.17. The percentage figures shown in column three of Table 4.17 do not sum to 100% because several officers weighted two or more factors equally, thereby causing a tie between factors ranked most important.

Table 4.17
Number of Officers Who Rated Each
Promotion Factor as Most Important

Promotion Factor	Number of Officers Who Ranked Factor First	Percent of Total Respondents*	an Relative Factor Weight for Officers Ranking Factor First
PME	119	37	.517
OER Ratings	137	43	.538
Assignment History	41	13	.413
Formal Education	61	19	.414
Aeronautical Rating	18	6	.431

* Percentages do not add to 100 because of ties between factor weightings.

In the interest of determining whether any one group of officers evaluated the promotability of the 32 hypothetical officers either excessively high or low, the mean value of all the decisions made by each officer group was computed. The mean evaluation score given by SOS officers was 4.42; for ACSC officers it was 4.61; and for AWC officers it was 4.81. Since the evaluation scale ranged from 1 (very low promotability)

to 7 (very high promotability), it appears that none of the officer groups evaluated the 32 hypothetical captains either excessively high or excessively low.

This concludes the summary of the research results. Further discussion and specific conclusions relating to these results are covered in the next chapter.

V Summary and Conclusions

Personnel evaluation and promotion is an important activity for any organization. Within the United States Air Force, the authority to promote commissioned officers is granted to the Secretary of the Air Force by public law. In addition, for the past several years other public laws have established grade limitations on the number of officers within specific grades or ranks. These grade limitations required that most officer promotions be based on the "best qualified" method of selection. However, the public law authorizing officer promotions does not specifically define the criteria to be used to select the best qualified officers. Futhermore, Air Force regulations also avoid clarifying the term best qualified by merely reflecting the intent of public law.

The concept of Air Force promotions stresses the whole person evaluation. However, specific criteria used in this evaluation are not clearly defined. Promotion boards are presented a "whole person concept" chart which lists general areas to evaluate. These criteria are not quantified or mandatory nor are any of them given precedence over any other. In general, one must conclude that the criteria used for selecting officers for promotion are based upon the judgment policies imposed by each promotion board. In other words, the lack of specific and clearly defined promotion criteria places the burden of determining what is important for promotion on each individual Air Force officer. Because individual judgment policies relating to promotion are apparently so important, this research effort investigates how Air Force officers render promotability decisions.

The research methodology used in this study is a logical follow-on to the work of many contemporary social scientiests who have used

judgment modeling techniques to capture individual policy and predict human judgment in a variety of real-world applications. A common term applied to judgment modeling techniques is "policy capturing."

Policy capturing analyzes decisions and yields a mathematical model of the cues on which decisions are based, weighting each according to the influence on the decision. The mathematical model used in this study was the linear regression model. Although some applications of the regression model are used to predict human judgment, this study utilized the model as an analytic instrument.

The policy capturing technique involves the use of controlled predictor variables as information cues for the experimental subject to employ as decision criteria in rendering a decision. The predictor variables used as cues in this study were five specific promotion factors which were perceived by the researcher to be operative in a decision relating to individual promotability. The promotion factors used in this study were: 1) Officer Effectiveness Report ratings, 2) professional military education, 3) formal education, 4) assignment history, and 5) aeronautical rating.

An experimental decision-making exercise was developed which required that a random sample of Air Force officers evaluate the promotability of 32 hypothetical captains. Decision criteria for the evaluations were based upon the five specific promotion factors. This decision exercise was administered to three separate groups of officers who were attending Air War College, Air Command and Staff College, and Squadron Officers' School. Policy capturing techniques were applied to the resulting data in order to analyze the manner in which the randomly selected officers rendered their promotability decisions.

Summary of the Research Results

The basic model used in this study was the linear regression model. This model was used in two separate analyses. The first analysis used regression to model the promotability decisions of each group of sampled officers. The second analysis used regression to model the decisions of each individual officer.

The Group Regression Analysis. There were two results from the group regression analysis that are pertinent to the study: 1) the significance of each promotion factor in the model and 2) the group squared multiple correlation coefficient (R^2).

The first result showed that all five promotion factors were utilized by each officer group in evaluating the promotability of the 32 hypothetical captains. This is a significant finding due to the fact that it indicates that officers actually utilize the whole person concept in judging the promotability of an individual.

The group squared multiple correlation coefficient values actually indicate the consistency of the group in making promotability decisions. The group R^2 value for SOS officers was .439; for ACSC officers the group R^2 was .382; and for AWC officers the R^2 value was .397. Since this value can range from 0 to 1, the results of the group regression R^2 values were relatively low. One might suspect that the low R^2 values may be the result of a greater number of random errors being introduced in the model because of such large groups. However, by looking at the mean R^2 values, derived from the individual regression model which were .762 for SOS officers, .776 for ACSC officers, and .791 for AWC officers, one can see that the individual officers were very consistent in making their promotability decisions. This large differential between the group R^2 values and

the mean R^2 values from the individual regressions implies that there is a wide variety in the way these officers made their promotability decisions.

Carrying this interpretation of the R^2 value a bit further, it is interesting to note that the SOS officers ($R^2 = .439$) had the least amount of variety in their decisions among the three groups. Likewise, the ACSC officers ($R^2 = .382$) had the most variety in the promotability decisions for the three groups.

F-Test of Significance on Group Regression Models. The F-test of significance performed on the group regression models showed that the three officer groups were statistically different in some respect regarding the importance placed on the five promotion factors. The specific differences were identified in another analysis using the Student t-test of significance.

Individual Regression Model Results. The individual regression model was used to determine the mean relative weight each officer group placed on the five promotion factors. In turn, the Student t-test of significance was used to determine where the significant differences existed among the three groups (the F-test performed on the group regression models already indicated there were differences).

This series of analyses found that all three groups of officers differed consistently on the relative weights placed on three of the five promotion factors: 1) PME, 2) OER ratings, and 3) formal education. This difference in factor weighting was most dramatic when comparing the SOS officers and the AWC officers. The SOS officers weighted PME and formal education higher than OER ratings while AWC officers weighted OER ratings over PME and formal education. These findings would

indicate that junior and senior ranking officers do not share a common perception regarding the importance of these promotion factors.

The ANOVA Model. The findings relating to the use of the ANOVA model confirm that each officer group utilized all five factors in their decisions. In addition, the ANOVA results showed that a linear model without interaction terms accounts for over 98% of the explained variance in the predictive power of the model.

Subjective and Objective Relative Weight Comparisons. A paired sample t-test was used to compare the subjective and objective relative weights placed on the promotion factors. The findings showed that SOS and ACSC officers did not apply their stated decision policies for four out of the five factors. In the case of AWC officers, it was found that these officers did apply their stated policies for four out of the five factors.

Conclusions and Implications of the Findings

In presenting the conclusions and implications drawn from the research findings, it may be beneficial to recall the five hypotheses proposed in Chapter I.

Hypothesis 1 was the primary hypothesis of the investigation.

> H1: Officers incorporate a whole person concept of evaluation in their promotability decisions by utilizing all the selected promotion factors in the promotability decision process.

The tests of significance associated with both the group regression model and the ANOVA model support hypothesis 1. The fact that the sampled officers utilized all five factors of promotion in making their decisions implies that Air Force officers are aware of the whole person concept and do apply the concept to promotability decisions.

The manner in which the three officer groups distributed the relative weights of the promotion factors provides another indication of the degree of influence the whole person concept had on the decisions. It is important to remember that the decision exercise was deliberately designed so that the performance ratings (OERs) of the 32 hypothetical captains were not materially different.

In cases where performance is approximately equal among eligible officers, one would expect that the decision-maker would apply the whole person evaluation in order to render the final promotability decision. The relative weights associated with the AWC senior officers indicate that these officers still considered performance (in the context of this study) very important, due to the fact that over 41% of the relative weight was applied to OER ratings. As a further indication of how strongly these officers felt about officer performance, one can inspect the results from the subjective versus the objective relative weights analysis. The mean subjective (perceived) relative weight for the OER factor was 43.1%. This subjective weight was not significantly different from the computed objective relative weight. The overall implication of these results indicates that the whole person concept of promotion evaluation influences the decisions of senior officers to a lesser extent than it does for the junior officers. Performance appears to be the most important criterion in the minds of senior officers even in cases where performance measures among eligible officers are very similar.

> H2: Junior officers place the same relative weights upon the criteria used for their promotability decisions as do senior officers.

The analysis performed on the promotion factor relative weights indicate that this hypothesis should be rejected. In the context of this

study, all three officer groups placed significantly different weights on OER ratings, PME, and formal education. Junior officers placed more weight on PME and formal education than on the OER ratings. In contrast, senior officers weighted the OER ratings more than either PME or formal education. This finding indicates disagreement among the officer groups concerning the importance placed upon three primary promotion factors.

This significant disagreement among the officer groups may imply some degree of dysfunctional behavior within the Air Force officer corps. There apparently is no common perception of what is important for promotion among the officer groups. In such a situation, junior officers may be directing efforts toward excelling in areas that their senior officers consider less important for promotion.

> H3: Air Force officers combine the promotability decision criteria in an essentially linear fashion to render a judgment.

The ANOVA analysis confirmed this hypothesis. A linear model without interaction terms adequately describes the manner in which officers combine the cues to render a decision. This finding may be useful for future research in modeling promotion decisions.

> H4: Individual officers accurately specify the relative weights they place upon the criteria used to render their promotability decisions.

From the analysis performed on the subjective and objective relative weights this hypothesis would be rejected. The results of the analysis showed that, as a group, the selected officers did not accurately specify their perceived judgment policy relating to their promotability decisions. However, the findings showed that the AWC officers (the most senior ranking officer group) did apply their subjective policies for four of the five promotion factors.

It is difficult to draw any specific conclusions from this finding, but it does have interesting implications. If Air Force officers do not accurately apply their perceived decision policy to a promotion decision, then in fact they may not accurately apply a specific Air Force promotion policy when directed to do so. This implication has some support in the form of a study conducted by Harrell (1975). This study found that when Air Force officers were given a specific policy to apply to their decisions, they attempted to utilize the policy. However, the officers did not accurately apply the policy in most cases.

> H5: Individual judgment policies are homogeneous within a group of officers, indicating that Air Force officers, as a group, make decisions relating to promotion in essentially the same manner.

Once again, the findings associated with this study show that this hypothesis should be rejected. The results from each officer group indicate that there is a wide variety in the decision-making of these officer groups. This would imply that there is a great deal of individuality prevalent within the Air Force officer corps when it comes to making promotion decisions.

Recommended Areas for Future Research

In most instances it is not wise to pronounce research findings from one single study as the solution to the problem. There are, of course, too many variables involved with most research studies to warrant such rash conclusions. For this reason it is recommended that policy capturing be applied, in a fashion similar to this study, to the promotability decision associated with Air Force colonel selections. Such a study would probably prove to be very interesting due to the fact that

colonel selections have become extremely competitive. In addition, such a study would either substantiate or refute the findings of this research effort.

Another area for research consideration relates to the finding in this study which implies that Air Force officers do not apply their subjective judgment policies when rendering promotability decisions. As this researcher envisions a study, an experiment could be designed which would incorporate a specific promotion policy to be applied to a promotion decision. Experimental subjects, randomly selected from the officer corps, would be asked to apply the stated policy to specific promotion decisions. The hypothesis for such a study would be that officers are capable of consistently applying a specific promotion policy to promotability decisions.

A final area for investigation holds some intriguing implications from the viewpoint of this researcher. As pointed out in the literature review chapter of this study, judgment models have proven to be at least as consistent as actual human judgments. In most cases, they have actually been more consistent than human judgments. The Air Force Academy study (Taylor and Wilsted, 1976), which was discussed in Chapter II, resulted in an application of judgment modeling to the cadet rating system for this very reason.

It appears that the present Air Force promotion system relies almost entirely on individual judgment policies for determining the criteria for promotion. This study indicates that there is a wide variety in the perceptions of individual officers regarding what is important for promotion. Such a situation might provide fertile ground for researching the applicability of a judgment model to the promotion process. This researcher

perceives an application similar to the Academy rating system. Selection boards would still be appointed, but their charge would be to evaluate each eligible officer on specific criteria. Air Force officials could then apply a specific policy to promotions by assigning weights to each criteria. The model would then be used to compute an overall evaluation for each eligible officer.

Bibliography

AFR 36-11. Permanent Promotion of Regular and Reserve Officers. Washington: Department of the Air Force, October 1974.

AFR 36-89. Temporary Promotion of Commissioned Officers. Washington: Department of the Air Force, January 1975.

Brunwsik, Egon. The Conceptual Framework of Psychology. Chicago: The University of Chicago Press, 1952.

Chow, Gregory C. "Tests of Equality Between Sets of Coefficients in Two Linear Regressions." Econometrica, 28:581-605 (July 1960).

Christal, Raymond E. "JAN: A Technique for Analyzing Group Judgment." The Journal of Experimental Education, 36:24-27 (1968).

Dawes, R. M. and B. Corrigan. "Linear Models in Decision Making." Psychological Bulletin, 81:95-106 (February 1974).

Ewing, Lee. "OER Scores Have Impact on O-4 Hikes." Air Force Times, 3 (January 17, 1977).

Ford, John N. "Officer Selection Boards and Due Process of Law." Military Law Review, 70:137-188 (Fall 1975).

Foster, George. "Best Qualified: What It Means." Air Force Times, 13 (January 17, 1973).

Freund, John E. Mathematical Statistics. Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1971.

Goldberg, L. R. "Five Models of Clinical Judgments: An Emperical Comparison Between Linear and Nonlinear Representation of the Human Inference Process." Organizational Behavior and Human Performance, 6:458-479 (1971).

Hammond, K. R. and D. A. Summers. "Cognitive Dependence on Linear and Nonlinear Cues." Psychological Review, 72:215-224 (1965).

Harrell, Adrian M. "Organizational Goals and the Decision-Making of Air Force Staff Officers: A Behavioral Experiment." AFIT Technical Report 75-6, Air Force Institute of Technology, Wright-Patterson Air Force Base, Ohio (May 1975).

Hays, W. L. Statistics for Social Scientists. New York: Holt, Rinehart and Winston, 1973.

Hoffman, Paul J. "The Paramorphic Representation of Clincial Judgment." Psychological Bulletin, 57:116-131 (March 1960).

Hoffman, Paul J., Paul Slovic, and Leonard Rorer. "An Analysis of Variance Model for Assessment of Cultural Cue Utilization in Clinical Judgment." Psychological Bulletin, 69:338-349 (May 1968).

Jones, Kirk M., Lawrence S. Mannis, Lester R. Martin, Jay L. Summers, and Gerald R. Wagner. "Judgment Modeling for Effective Policy and Decision Making." Technical Report sponsored by the Air Force Office of Scientific Research, Air Force Systems Command. Defense Documentation Center Number AD-A033186 (December 1976).

Kleinmuntz, Benjamin (editor). Formal Representation of Human Judgment. New York: John Wiley and Sons, Inc., 1968.

Letter of Instructions (Central Temporary Major Board, Calendar Year 1976). Headquarters Air Force Military Personnel Center, Randolph Air Force Base, Texas.

Nie, H. H., C. H. Hull, J. G. Jenkins, K. Steinbrenner, and D. H. Bent. Statistical Package for the Social Sciences. New York: McGraw-Hill Book Company, 1975.

Meyer, John C. "Officer Promotions." Combat Crew, 33:1-2 (September 1973).

O'Berry, Carl G. "An Experimental Test for Goal Congruence in an Air Force Major Command." Unpublished masters thesis. Air Force Institute of Technology, Wright-Patterson Air Force Base, Ohio (September 1977).

"Short Sheeted Again." Armed Forces Journal International, 110:14 (January 1973).

Slovic, Paul, Baruch Fischoff, and Sarah Licktenstein. "Behavioral Decision Theory." Technical Report sponsored by Defense Advanced Research Projects Agency. Defense Documentation Center Number AD-A036744 (September 1976).

Slovic, Paul, Dan Fleissner, and W. Scott Bauman. "Analyzing the Use of Information in Investment Decision-Making: A Methodological Proposal." Journal of Business, 45:283-301 (April 1972).

Slovic, Paul and Sarah Licktenstein. "Comparison of Bayesian and Regression Approaches to the Study of Information Processing in Judgment." Organizational Behavior and Human Performance, 6:649-744, 1971.

Smith, Thomas. "A Flow Chart for Policy Capturing." Report Number 147. Institute for Behavioral Science, University of Colorado, 1972.

Taylor, Robert L. "Capturing Judgment Policies: An Application for Cadet Rating at the United States Air Force Academy." United States Air Force Academy Research Report 73-7. United States Air Force Academy, Colorado, July 1973.

Taylor, Robert L. and William D. Wilsted. "Capturing Judgment Policies in Performance Rating." Industrial Relations, 15:216-224 (May 1976).

United States Code. Title 10--Armed Forces. Washington, D.C. 1970 edition.

Related Sources

Adelman, Leonard, Thomas R. Stewart, and K. R. Hammond. "A Case History of the Application of Social Judgment Theory to Policy Formulation." Policy Sciences, 6:137-159 (1975).

Beach, Lee Roy. "Multiple Regression as a Model for Human Information Utilization." Organizational Behavior and Human Performance, 2:276-289 (August 1967).

Beusse, William E. "Attitudes of Military Officers Toward Promotion." Technical Report. Personnel Research Division, Lackland Air Force Base, Texas. Defense Documentation Center Number AD-A036-108 (February 1976).

Castillan, N. "Comments on the 'Lens Model' Equation and the Analysis of Multiple-Cue Judgment Tasks." Psychometrika, 38:87-100 (March 1973).

Christal, Raymond E. "Selecting A Harem and Other Applications of the Policy Capturing Model." The Journal of Experimental Education, 36:35-41 (Summer 1968).

Darlington, Richard B. "Multiple Regression in Psychological Research and Practice." Psychological Bulletin, 69:161-180 (March 1968).

Goldberg, L. R. "Man Versus Model of Man. A Rationale, Plus Some Evidence for a Method of Improving on Clinical Inferences." Psychological Bulletin, 73:422-432 (June 1970).

Goldberg, Lewis R. "Simple Models or Simple Processes? Some Research on Clinical Judgments." American Psychologist, 23:483-486 (July 1968).

Hursch, C. J., K. R. Hammond, and J. Hursh. "Some Methodological Consideration in Multiple-Cue Probability Studies." Psychological Review, 71:42-60 (1964).

Schenk, E. A. and J. C. Naylor. "A Cautionary Note Concerning the Use of Regression Analysis forCapturing Strategies of People." Educational and Psychological Measurement, 28:3-7 (1968).

Slovic, Paul. "Cue Consistency and Cue Utilization in Judgment." American Journal of Psychology, 79:427-434 (1966).

APPENDIX A

DECISION-MAKING EXERCISE

A DECISION-MAKING EXERCISE

FOR

AIR FORCE OFFICERS

THIS IS NOT A QUESTIONNAIRE. It is a decision-making exercise that is designed to investigate how individuals such as yourself arrive at their decisions. The entire exercise will take about fifteen minutes to complete. Your participation in this investigation will be kept strictly confidential. You will not be identified in the final report (a masters thesis at the Air Force Institute of Technology). Please cooperate so that we may learn more about how Air Force officers such as yourself arrive at their decisions.

If you would like to receive information about your decision-making and how it compares with that of your contemporaries, please print your name and address in the space provided below. A summary comparison of your results with those of others will be mailed to you, in confidence, after completion of the study (December 1977).

____
Name

____
Address

____
City State Zip Code

USAF SCN 77-156
(Expires 31 Dec 77)

PRIVACY STATEMENT

In accordance with AFR 12-35, paragraph 30, the following information is provided as required by the Privacy Act of 1974:

a. Authority

(1) 5 U.S.C. 301, Departmental Regulations: and/or

(2) 10 U.S.C. 80-12, Secretary of the Air Force, Powers and Duties, Delegation By.

b. Principal purposes. The survey is being conducted to collect information to be used in research aimed at illuminating and providing inputs to the solution of problems of interest to the Air Force and/or DOD.

c. Routine uses. The survey data will be converted to information for use in research of management related problems. Results of the research, based on the data provided, will be included in a written masters thesis and may also be included in published articles, reports, or texts. Distribution of the results of the research, based on the survey data, whether in written form or orally presented, will be unlimited.

d. Participation in this survey is entirely voluntary.

e. No adverse action of any kind may be taken against any individual who elects not to participate in any or all of this survey.

BACKGROUND INFORMATION

Please circle the most appropriate answer for each of the following questions:

1. What is your present rank?

    a. 2nd Lt   d. Major   g. Other (specify) ___________

    b. 1st Lt   e. Lt Col

    c. Captain   f. Colonel

2. What aeronautical rating do you hold?

    a. Pilot   c. Nonrated

    b. Navigator   d. Other (specify) ___________

3. How would you classify your present or most recent nonschool assignment?

    a. Squadron   d. Numbered AF Hq   g. Other (specify) _________

    b. Wing   e. Major Air Comd Hq

    c. Division Hq   f. USAF Hq

4. What is the highest level of assignment you have had during your Air Force career?

    a. Squadron   d. Numbered AF Hq   g. Other (specify) _________

    b. Wing   e. Major Air Comd Hq

    c. Division Hq   f. USAF Hq

5. What level of formal education have you completed?

    a. Undergraduate degree   c. Doctoral degree

    b. Masters degree   d. Other (specify) ____________

6. Indicate the largest number of people you have ever had under your supervision or command.

    a. Less than 10   d. At least 30 but less than 40

    b. At least 10 but less than 20   e. At least 40 but less than 50

    c. At least 20 but less than 30   f. At least 50 but less than 100

    g. 100 or more

Indicate the overall rating you received on your last three (or less) controlled Officer Effectiveness Reports (OERs). Use ratings of "1," "2," "3," etc.

7. Most recent OER: _______

8. Second most recent OER: _______

9. Third most recent OER: _______

INSTRUCTIONS FOR THE EXERCISE

On the following pages you are asked to judge the promotability of a number of hypothetical Air Force captains who are in the primary zone for promotion to the rank of major. The individual cases presented to you are not intended to be representative of Air Force captains in general. You should consider these individuals to be members of a very special group of Air Force captains who are exactly alike in all respects except for five key areas. These five key areas are not necessarily intended to represent the decision criteria that might be used by an Air Force promotion board. Above all, this exercise is not an attempt to represent the promotion board process. You are to evaluate each individual captain based solely upon the information provided for you and your perception of what is important for promotability within the current Air Force environment. The five key areas for your decision criteria are defined as follows:

1. OER RATINGS - The Air Force Chief of Staff announced a new policy on controlled OER ratings in August 1977. You are asked to project the implications of this new policy forward into the future and to assume this new policy has been in effect for all the OER ratings presented in this exercise. This means, for the OERs shown, that 22% of the officers could receive an overall rating of "1" while 78% could receive a "2" or lower. The three most recent OER ratings for each individual, all received under the new policy, are presented to you in order from earliest to most recent. Only two rating patterns are presented, so each individual will have either one or the other, as follows:

   First Pattern: 2, 1, 1
   Second Pattern: 2, 2, 1

   Consider that all of the officers involved received exactly the same OER ratings for all previous reports.

2. ASSIGNMENT HISTORY - A term used to indicate whether the individual has completed a headquarters-level assignment (division or higher) or has served entirely at the base level (squadron, wing, etc.).

3. PME HISTORY - Professional Military Education (PME) history indicates whether the individual has completed Squadron Officers' School by either correspondence or in resident.

4. AERONAUTICAL RATING - For this exercise, aeronautical rating distinguishes between pilots and nonrated personnel.

5. FORMAL EDUCATION - The level of formal schooling an individual has completed. For this exercise, this will be either a bachelors degree or a masters degree.

Each case is presented to you in the following format:

CAPTAIN #0

DECISION CRITERIA	INDIVIDUAL PERFORMANCE
FORMAL EDUCATION	
ASSIGNMENT HISTORY	
AERONAUTICAL RATING	(specific information will be given)
OER RATINGS	
PME HISTORY	

DECISION #0

1 . . . . 2 . . . . 3 . . . . 4 . . . . 5 . . . . 6 7

Low Promotability — Moderate Promotability — High Promotability — Very High Promotability

You are to circle the number that reflects your decision concerning the promotability of the individual. You may make your decisions in whatever manner you wish. You should not, however, change a decision after you have completed your response.

PLEASE COMPLETE ALL CASES, AS EACH INDIVIDUAL CASE IS DIFFERENT

CAPTAIN #1

DECISION CRITERIA	INDIVIDUAL PERFORMANCE
PME HISTORY	Squadron Officers' School
ASSIGNMENT HISTORY	Entirely at base level
AERONAUTICAL RATING	Pilot
OER RATINGS	2, 1, 1
FORMAL EDUCATION	Masters degree

DECISION #1

1 . . . . 2 . . . . 3 . . . . 4 . . . . 5 . . . . 6 7

Low Promotability — Moderate Promotability — High Promotability — Very High Promotability

CAPTAIN #2

DECISION CRITERIA	INDIVIDUAL PERFORMANCE
AERONAUTICAL RATING	Pilot
FORMAL EDUCATION	Masters degree
ASSIGNMENT HISTORY	Headquarters tour completed
PME HISTORY	None
OER RATINGS	2, 2, 1

DECISION #2

1 . . . . 2 . . . . 3 . . . . 4 . . . . 5 . . . . 6 7

Low Promotability — Moderate Promotability — High Promotability — Very High Promotability

NOTE: Pages 8 through 22 of the decision exercise have been omitted from this appendix. These pages contained the other combinations of the five cues and, therefore, represented hypothetical captains numbers 3 through 32.

LAST TASK

Please indicate the relative importance that you believe you placed upon each of the five criteria during the exercise by distributing 100 points to these criteria. The most important criterion should receive the most points, etc.

CRITERIA	ASSIGNED POINTS
AERONAUTICAL RATING	________
OER RATINGS	________
ASSIGNMENT HISTORY	________
PME HISTORY	________
FORMAL EDUCATION	________
Total Points:	100

THANKS FOR YOUR COOPERATION. REMEMBER! SHOULD YOU DESIRE TO HAVE YOUR INDIVIDUAL RESULTS MAILED TO YOU, PLEASE FILL IN YOUR NAME AND ADDRESS IN THE SPACE PROVIDED ON PAGE 1.

APPENDIX B

F-TEST COMPUTATIONAL FORM

APPENDIX B

Calculation of F-Test Values

The F-test values used to compare regression models in this study were calculated using the following formula:

$$F_o = \frac{\left[SS_e - \sum_{j=1}^{p} SS_{ej}\right] / \left[(p-1)(K+1)\right]}{\left[\sum_{j=1}^{p} SS_{ej}\right] / \left[n-p(k+1)\right]}$$

where SS_e is the residual sum of squares derived by regressing all compared groups of decisions together, SS_{ej} is the residual sum of squares for the jth group of evaluations, p is the number of groups being compared (number of subsets of data in the regression), k is the number of predictor variables (four, in all cases for this study), and n is the total number of decisions in all groups being compared.

The null hypothesis being tested is

$$H_o: \underline{\beta}_1 = \underline{\beta}_2 = \underline{\beta}_3 = \cdots = \underline{\beta}_p, \text{ where } \underline{\beta}_i = \begin{matrix} \beta_o \\ \beta_1 \\ \vdots \\ \beta_k \end{matrix}$$

The alternate hypothesis is

$$H: \underline{\beta}_i \neq \underline{\beta}_j, \text{ for at least one i, j pair.}$$

The null hypothesis is rejected if

$$F_o > F_{\alpha}, \left[(p-1)(k+1)\right], \left[n-p(k+1)\right]$$

where $\alpha = 0.05$ in all comparisons made for this study.

APPENDIX C

EXAMPLE OF FEEDBACK PROVIDED

TO EXERCISE RESPONDENTS

SCHOOL OF ENGINEERING
AIR FORCE INSTITUTE OF TECHNOLOGY
WRIGHT-PATTERSON AIR FORCE BASE
OHIO

12 DECEMBER 1977

TO:

THANK YOU FOR PARTICIPATING IN THE RECENT DECISION ANALYSIS EXERCISE. YOU REQUESTED AN ANALYSIS OF YOUR PERFORMANCE IN THAT EXERCISE. THE FOLLOWING STATISTICS COMPARE YOUR 32 PROMOTABILITY DECISIONS AGAINST THOSE MADE BY OTHERS IN YOUR PME CLASS AS A GROUP. THE WEIGHT SHOWN FOR EACH PROMOTION FACTOR IS THE RELATIVE IMPORTANCE GIVEN THAT FACTOR IN DECIDING THE PROMOTABILITY OF THE HYPOTHETICAL CAPTAINS.

	WEIGHTS YOU USED IN EVALUATING THE 32 HYPOTHETICAL CAPTAINS:	WEIGHTS ASSIGNED BY OTHERS IN YOUR PME CLASS AS A GROUP:
PME HISTORY	40.8%	34.9%
ASSIGNMENT HISTORY	25.4%	12.8%
AERONAUTICAL RATING	5.5%	8.7%
OER RATINGS	2.8%	20.8%
FORMAL EDUCATION	25.4%	23.6%

THE CONSISTENCY WITH WHICH YOU APPLIED THE WEIGHTS INDICATED ABOVE IN EVALUATING THE CAPTAINS WAS 85.2%. THE CONSISTENCY DISPLAYED BY YOUR FELLOW CLASSMATES AS A GROUP WAS 76.2%. THIS MEANS THAT YOUR PERSONAL EVALUATION POLICY, ONCE ESTABLISHED, WAS EFFECTIVELY APPLIED IN 85.2% OF THE 32 DECISIONS YOU MADE IN THE EXERCISE.

THANKS AGAIN FOR YOUR COOPERATION IN THIS RESEARCH EFFORT.

CAPTAIN PHILIP E. GLENN
AFIT/ENS
WRIGHT-PATTERSON AFB, OH 45433

Vita

Captain Philip E. Glenn [redacted] After graduation from Fulton High School in 1964, he attended one year of undergraduate school at Westminster College located in Fulton, Missouri. The following year Captain Glenn transferred to the University of Missouri at Columbia, Missouri. While at the University of Missouri, he majored in electrical engineering and physics. He was a member of Eta Kappa Nu, the national electrical engineering honorary fraternity, and he served in several offices in the local chapter of the Institute for Electrical and Electronic Engineers (IEEE). In 1969 Captain Glenn garduated with bachelor of science degrees in electrical engineering and physics.

Captain Glenn entered the United States Air Force in June, 1970 after receiving his commission through the Reserve Officer Training Corps (ROTC) program. His first assignment was pilot training at Vance Air Force Base, Oklahoma. Following pilot training, Captain Glenn spent the next five years flying the HC-130 aircraft under the command of the Aerospace Rescue and Recovery Service. Assignments following Vance Air Force Base included Yokota Air Base, Tokyo, Japan, Kadena Air Base, Okinawa, Japan, and Eglin Air Force Base, Florida. In August 1976 Captain Glenn reported to the Air Force Institute of Technology School of Engineering to work toward a Master of Science degree in Systems Management. Following graduation, Captain Glenn will be assigned as a Communications-Electronics Engineer with the Air Force Communications Service.

[redacted]

[redacted]

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AFIT/GSM/SM-77D-20

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle): A JUDGMENT ANALYSIS APPROACH TO EXAMINING A "WHOLE PERSON" CONCEPT OF AIR FORCE PROMOTIONS

5. TYPE OF REPORT & PERIOD COVERED: MS

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s): Philip E. Glenn, Captain USAF

8. CONTRACT OR GRANT NUMBER(s)

9. PERFORMING ORGANIZATION NAME AND ADDRESS: Air Force Institute of Technology (AFIT-EN), Wright-Patterson AFB, Ohio 45433

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: N/A

11. CONTROLLING OFFICE NAME AND ADDRESS: N/A

12. REPORT DATE: December 1977

13. NUMBER OF PAGES: 93

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)

15. SECURITY CLASS. (of this report): Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

Approved for public release; IAW AFR 190-17

JERRAL F. GUESS, Captain, USAF
Director of Information

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

officer promotions, promotion, judgment analysis, decision-making, policy capturing, personnel evaluation

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The United States Air Force stresses the "whole person" concept in the evaluation of all commissioned officers who are eligible for promotion. However, no specific policy has been established by the Air Force for applying this whole person evaluation. The separate promotion boards are presented a whole person chart which recommends factors and areas for evaluation. The criteria depicted on this chart, however, are not mandatory nor do they have precedence over any other criteria. It appears, then, that each promotion board establishes the policy which is utilized in selecting commissioned —

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

officers for promotion.

This research effort investigates the extent to which a whole person concept of promotion evaluation influences the beliefs held by Air Force officers concerning the promotability of captains to the rank of major. In order to do so, a judgment analysis approach known as policy capturing was utilized.

A random sample of commissioned officers attending Squadron Officers' School, Air Command and Staff College, and Air War College was administered a decision-making exercise. In this exercise, the selected officers were asked to evaluate the promotability of 32 hypothetical captains based upon five key promotion factors which were used as decision criteria. The five promotion factors used in the research were: 1) Officer Effectiveness Report ratings, 2) professional military education, 3) formal education, 4) assignment history, and 5) aeronautical rating. Data collected from this decision exercise were used to test specific hypotheses concerning the decision-making behavior of the selected officers.

The findings of the research indicate that Air Force officers are aware of the whole person concept of promotion evaluation and utilize it in their promotability decisions. Junior officers, however, placed significantly different weights on three of the five factors when compared with the weights placed on the same factors by senior officers. The data indicates that officers combine the decision criteria in an essentially linear fashion when rendering their promotability decisions. Finally, the findings showed that the individual decision-making policies among the officers were not homogeneous, which indicates that there is a wide variety in the way Air Force officers make promotability decisions.

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)